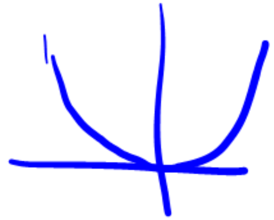
Let $X \sim N(0,1)$ & $Y = X^2$. 

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad -\infty < x < \infty$$

$x = -\sqrt{y}$, $\qquad\qquad x = \sqrt{y}$

$$\frac{dx}{dy} = -\frac{1}{2\sqrt{y}} \qquad\qquad \frac{dx}{dy} = \frac{1}{2\sqrt{y}}$$

$\left|\frac{dx}{dy}\right| = \frac{1}{2\sqrt{y}}$ in both parts

So the pdf of Y is then

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y/2} \cdot \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-y/2} \cdot \frac{1}{2\sqrt{y}}$$

$$= \begin{cases} \dfrac{1}{\sqrt{2\pi y}} e^{-y/2}, & y > 0 \\ 0, & y \leq 0 \end{cases} \longrightarrow \frac{1}{2^{1/2}\Gamma\frac{1}{2}} y^{\frac{1}{2}-1} e^{-y/2}$$

$$\text{Gamma}\left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

Jointly Distributed Random Variables

$X_1 \to$ body temperature

$X_2 \to$ Oxygen level

$X_3 \to$ Systolic BP

$X_4 \to$ Dystolic BP

$X_5 \to$ Pulse Rate

$$\underline{X} = (X_1, X_2, X_3, X_4, X_5)$$

$X_1 \to$ marks in Course 1

$X_2 \to$ marks in Course 2

$\cdots$

$X_6 \to$ marks in Course 6

$$\underline{X} = (X_1, \ldots, X_6)$$

$\underline{X} = (X_1, \ldots, X_k) : \Omega \to \mathbb{R}^k$

$\to$ k-dimensional r. vector

cdf of a Random Vector

$F_{\underline{X}}(\underline{x}) = P(X_1 \leq x_1, \ldots, X_k \leq x_k)$.

Let us first take the case $k = 2$

$(X, Y) \in \mathcal{X} \times \mathcal{Y}$

$F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$

$\forall (x, y) \in \mathbb{R}^2$

<u>Properties</u>: 1. $\lim_{x \to -\infty} F_{X,Y}(x,y) = 0$

2. $\lim_{y \to -\infty} F_{X,Y}(x,y) = 0$

3. $\lim_{x \to +\infty} F_{X,Y}(x,y) = F_Y(y)$

4. $\lim_{y \to +\infty} F_{X,Y}(x,y) = F_X(x)$

5. $F(\cdot,\cdot)$ is nondecreasing in each of its arguments

6. $F(\cdot,\cdot)$ is continuous from right in each of its arguments.

Discrete Case: Suppose both X and Y are discrete. Then the joint pmf $p_{X,Y}(x_i, y_j)$ is defined as

(i) $P(X = x_i, Y = y_j) = p_{X,Y}(x_i, y_j)$

(ii) $p_{X,Y}(x_i, y_j) \geq 0$

(iii) $\sum \sum p_{X,Y}(x_i, y_j) = 1.$

Example: 10 computers in a showroom

$\swarrow$ 5 (good)

$\rightarrow$ 2 (defective monitor)

$\searrow$ 3 (defective CPU)

Suppose 2 computers are selected at random

$X \rightarrow$ no of computers with DM

$Y \rightarrow$ no of computers with DC

We want joint pmf of (X,Y)

$$p_{X,Y}(0,0) = P(X=0, Y=0) = \frac{\binom{5}{2}}{\binom{10}{2}} = \frac{10}{45} = \frac{2}{9}$$

$$p_{X,Y}(0,1) = \frac{\binom{5}{1}\binom{3}{1}}{\binom{10}{2}} = \frac{15}{45} = \frac{1}{3}$$

The pmf can be represented in the tabular form

$X \backslash Y$	0	1	2	$p_X(\cdot)$
0	$10/45$	$15/45$	$3/45$	$28/45$
1	$10/45$	$6/45$	0	$16/45$
2	$1/45$	0	0	$1/45$
$p_Y(\cdot)$	$21/45$	$21/45$	$3/45$	

$p_X(\cdot) \rightarrow$ marginal pmf of X

marginal pmf of Y

$$P(X \leq 1, Y \leq 1) = p_{X,Y}(0,0) + p_{X,Y}(0,1) + p_{X,Y}(1,0) + p_{X,Y}(1,1)$$

$$= \frac{10}{45} + \frac{15}{45} + \frac{10}{45} + \frac{6}{45} = \frac{41}{45}$$

Marginal pmf of X is

$$p_X(x_i) = \sum_{y_j \in \mathcal{Y}} p_{X,Y}(x_i, y_j)$$

The marginal pmf of Y is

$$p_Y(y_j) = \sum_{x_i \in \mathcal{X}} p_{X,Y}(x_i, y_j)$$

The conditional pmf of X given $Y = y_j$

$$p_{X|Y=y_j}(x_i | y_j) = P(X = x_i | Y = y_j)$$

$$= \frac{p_{X,Y}(x_i, y_j)}{p_Y(y_j)}, \quad x_i \in \mathcal{X}$$

Similarly the conditional pmf of Y given $X = x_i$

$$p_{Y|X=x_i}(y_j|x_i) = P(Y=y_j \mid X=x_i)$$

$$= \frac{p_{X,Y}(x_i, y_j)}{p_X(x_i)}, \quad y_j \in \mathcal{Y}.$$

<u>Example</u>: Find the conditional pmf of Y given $X=0$.

$$p_{Y|X=0}(0|0) = \frac{p_{X,Y}(0,0)}{p_X(0)} = \frac{10/45}{28/45} = \frac{10}{28}$$

$$p_{Y|X=0}(1|0) = \frac{p_{X,Y}(0,1)}{p_X(0)} = \frac{15/45}{28/45} = \frac{15}{28}$$

$$p_{Y|X=0}(2|0) = \frac{p_{X,Y}(0,2)}{p_X(0)} = \frac{3/45}{28/45} = \frac{3}{28}.$$

Let (X,Y) be jointly distributed continuous random variables with joint pdf $f_{X,Y}(x,y)$

(i) $f_{X,Y}(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2$

(ii) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\,dx$$

(iii) For any measurable set $B \subset \mathbb{R}^2$,

$$P((X,Y) \in B) = \iint_B f_{X,Y}(x,y)\,dx\,dy = \iint_B f_{X,Y}(x,y)\,dy\,dx$$

The marginal pdf of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

and the marginal pdf of Y is

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

Conditional pdf of X given $Y=y$ is

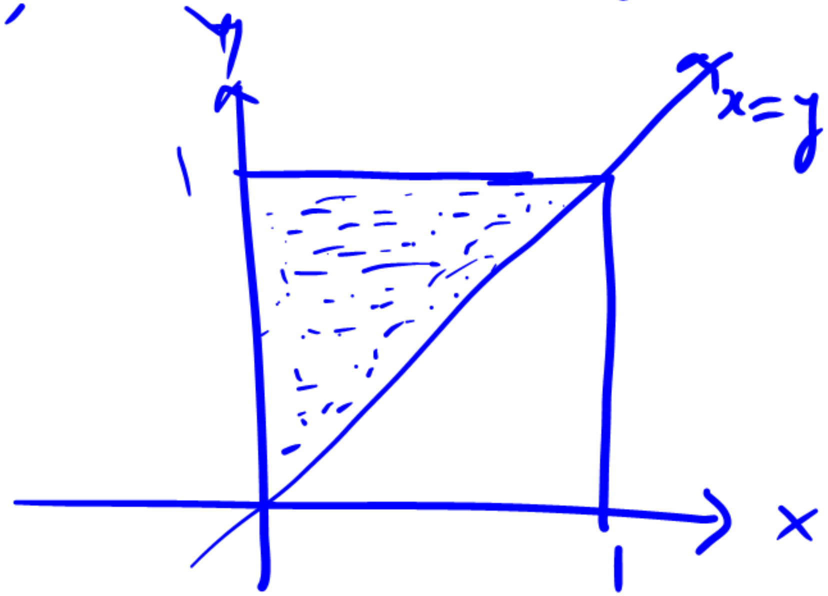
$$f_{X|Y=y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Conditional pdf of Y given $X=x$ is

$$f_{Y|X=x}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

<u>Example</u>: Let (X,Y) be jointly continuous

with pdf $f_{X,Y}(x,y) = \begin{cases} 10xy^2, & 0<x<y<1 \\ 0, & \text{otherwise} \end{cases}$



The marginal pdf of X is

$$f_X(x) = \int_x^1 10xy^2\,dy$$

$$= \begin{cases} \frac{10}{3}x(1-x^3), & 0<x<1 \\ 0, & \text{ew} \end{cases}$$

The marginal pdf of Y is

$$f_Y(y) = \int_0^y 10xy^2\,dx = \begin{cases} 5y^4, & 0<y<1 \\ 0, & ew \end{cases}$$

$$P(X<\tfrac{1}{4}) = \int_0^{1/4} f_X(x)\,dx = \int_0^{1/4} \frac{10}{3}x(1-x^3)\,dx$$

$$= \frac{10}{3}\left[\frac{1}{32} - \frac{1}{5\cdot 4^5}\right] = \ldots\ldots$$

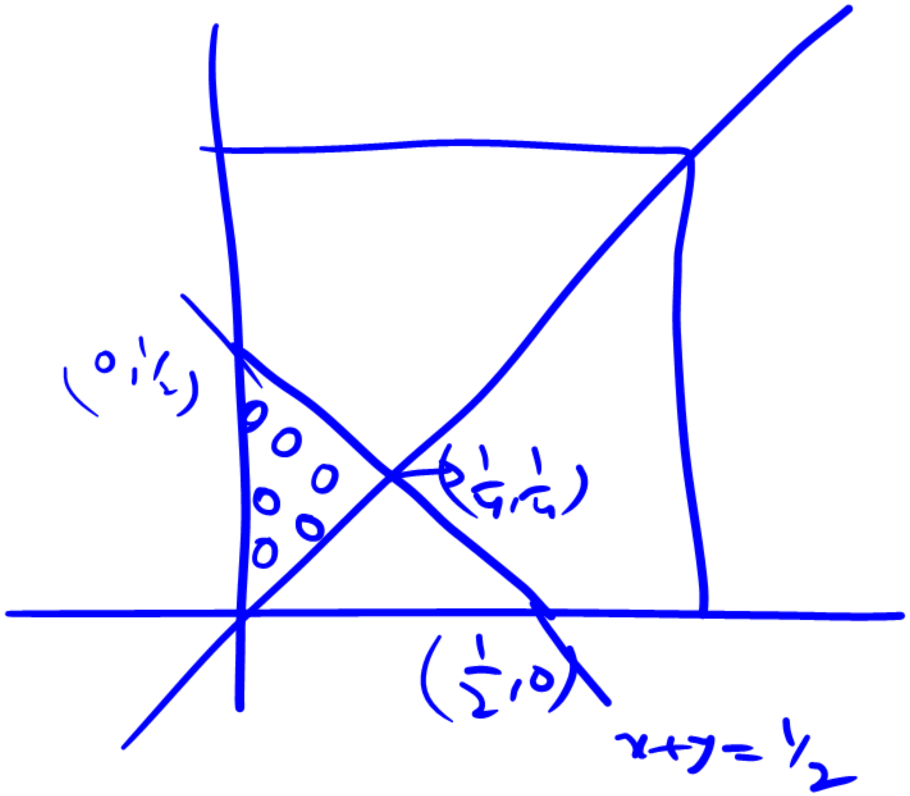
$$P(Y>3/4) = \int_{3/4}^{1} 5y^4\,dy = \left[1-\left(\frac{3}{4}\right)^5\right]$$

$$P\left(0 < X+Y < \frac{1}{2}\right)$$

$$= \iint_{0<x+y<\frac{1}{2}} 10\,x y^2 \, dy\, dx$$

$$= \int_0^{1/4} \int_x^{\frac{1}{2}-x} \left(10 x y^2\right) dy\, dx$$

$$= \frac{10}{3} \int_0^{1/4} x\left[\left(\frac{1}{2}-x\right)^3 - x^3\right] dx = \ldots$$

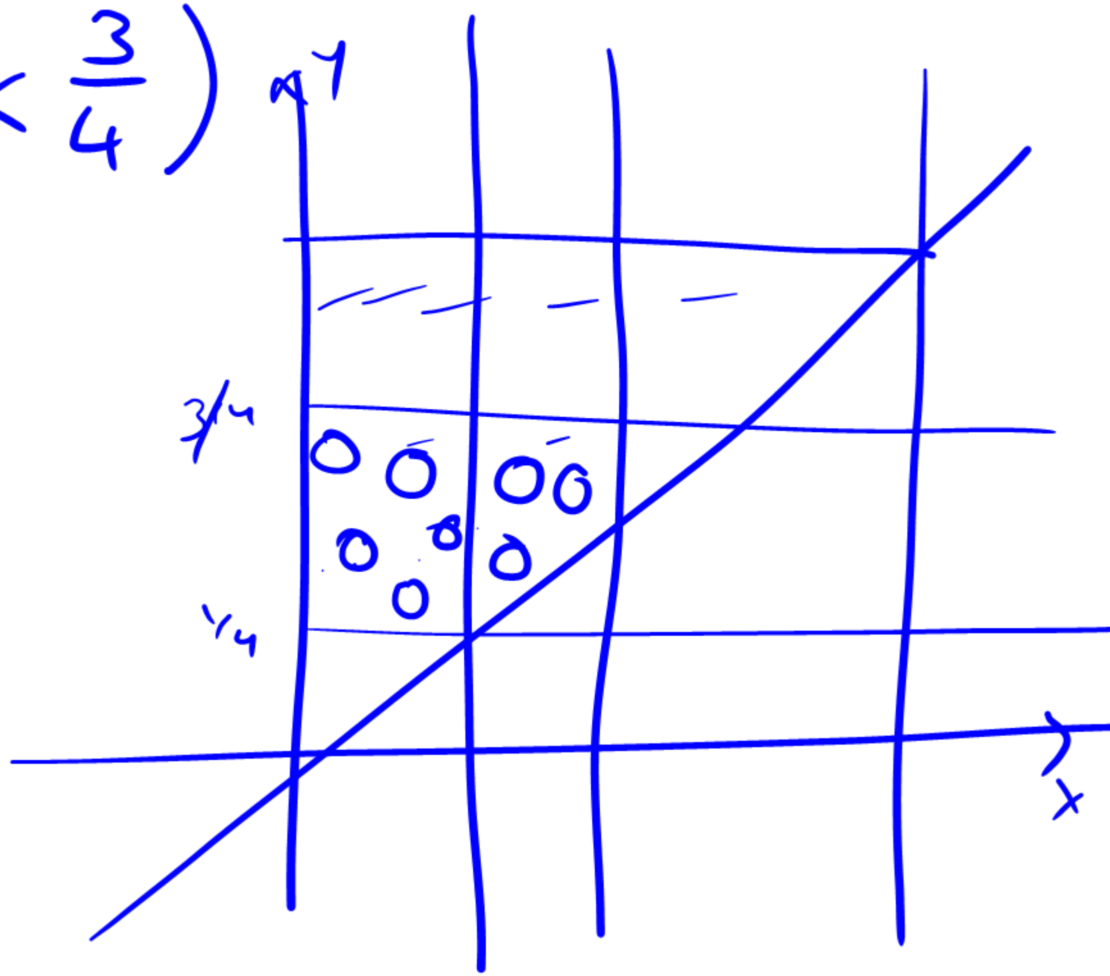


$$P\left(0 < X < \frac{1}{2},\ \frac{1}{4} < Y < \frac{3}{4}\right)$$

$$= \int_{1/4}^{3/4} \int_{0}^{1/4} 10xy^2 \, dx\, dy$$

$$+ \int_{1/4}^{1/2} \int_{x}^{3/4} 10xy^2 \, dy\, dx$$

$$= \cdots$$



The conditional pdf of X given $Y = y$ is

$$f_{X|Y=y}(x|y) = \begin{cases} \dfrac{2x}{y^2}, & 0<x<y, \quad 0<y<1 \\ 0, & ew \end{cases}$$

The conditional pdf of Y given $X=x$ is

$$f_{Y|X=x}(y|x) = \begin{cases} \dfrac{3y^2}{1-x^3}, & x<y<1, \quad 0<x<1 \\ 0, & ew \end{cases}$$

$$P(X<1/2 \mid Y=3/4) = \int_0^{1/2} \frac{32}{9}x\,dx = \frac{4}{9}$$

$$f_{X|Y=3/4}(x) = \begin{cases} \frac{32}{9}x, & 0<x<3/4 \\ 0, & ew \end{cases}$$

$$P(Y<1/2 \mid X=1/4)$$

$$f_{Y|X=\frac{1}{4}}(y) = \begin{cases} \frac{64}{21}y^2, & \frac{1}{4}<y<1 \\ 0, & ew \end{cases}$$

$$P(Y<\tfrac{1}{2} \mid X=1/4) = \int_{1/4}^{1/2} \frac{64}{21}y^2\,dy = \frac{1}{9}$$

Independence of Random Variables

We say that random variables X and Y are independently distributed if

$$F_{X,Y}(x,y) = F_X(x)\,F_Y(y) \quad \forall (x,y) \in \mathbb{R}^2$$

In case X, Y are discrete, the condition can be written as

$$p_{X,Y}(x_i, y_j) = p_X(x_i)\,p_Y(y_j) \quad \forall (x_i, y_j)$$

In case X, Y are continuous, the condition can be written as

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad \forall (x,y) \in \mathbb{R}^2$$

<u>Examples</u>: $p_{X,Y}(0,0) = \frac{1}{4}$, $p_{X,Y}(1,0) = \frac{1}{4}$

$p_{X,Y}(0,1) = \frac{1}{4}$, $p_{X,Y}(1,1) = \frac{1}{4}$

Then $p_X(0) = \frac{1}{2}$, $p_X(1) = \frac{1}{2}$, $p_Y(0) = \frac{1}{2}$, $p_Y(1) = \frac{1}{2}$

So $p_{X,Y}(x_i, y_j) = p_X(x_i)\, p_Y(y_j) \quad \forall (x_i, y_j)$

So X and Y are independently distributed.

2. $f_{X,Y}(x,y) = \begin{cases} e^{-2y}, & 0<x<2,\ y>0 \\ 0, & \text{ew} \end{cases}$

$$f_X(x) = \int_0^\infty e^{-2y}\,dy = \frac{1}{2}, \quad 0<x<2$$

$$f_Y(y) = \int_0^2 e^{-2y}\,dx = 2e^{-2y}, \quad y>0$$

So $f_{X,Y}(x,y) = f_X(x) f_Y(y) \ \forall (x,y)$

So X and Y are independently distributed.

Joint Expectation

$$g(x, y)$$

$$E\, g(X, Y) = \sum_{(x_i, y_j)} \sum g(x_i, y_j)\, p_{X,Y}(x_i, y_j) \quad \text{discrete}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)\, f_{X,Y}(x, y)\, dx\, dy \Longleftrightarrow dy\, dx \quad \text{continuous}$$

provided the series/integral on the right are absolutely convergent.

Product Moments:

$$\mu'_{r,s} = E(X^r Y^s)$$

$\rightarrow (r,s)^{th}$ noncentral product moment

Special Cases $\mu'_{1,0} = E(X) = \mu_X$

$$\mu'_{0,1} = E(Y) = \mu_Y$$

$$\mu'_{1,1} = E(XY) = \mu_{XY}$$

$$\mu_{r,s} = E\left[(X-\mu_X)^r (Y-\mu_Y)^s\right]$$

$\rightarrow (r,s)^{th}$ central product moment

$$\mu_{1,1} = E\left[(X-\mu_X)(Y-\mu_Y)\right] = \text{Covariance}(X,Y)$$

$$= E\left(XY - X\mu_Y - \mu_X Y + \mu_X \mu_Y\right)$$

$$= E(XY) - \mu_X\mu_Y - \cancel{\mu_X\mu_Y} + \cancel{\mu_X\mu_Y}$$

$$= \mu_{XY} - \mu_X\mu_Y = E(XY) - E(X)E(Y)$$

Let r.v.'s X and Y be independent. Suppose they are continuous with densities $f_X(x)$ & $f_Y(y)$ respectively.

$$E(X^r Y^s) = \int\int x^r y^s f_{X,Y}(x,y)\,dx\,dy$$

$$= \int\int x^r y^s f_X(x) f_Y(y)\,dx\,dy$$

$$= \left(\int x^r f_X(x)\,dx\right)\left(\int y^s f_Y(y)\,dy\right)$$

$$= E(X^r)\ E(Y^s).$$

In particular, if X and Y are independent then $\mathrm{Cov}(X,Y) = 0$

We also use notations

$\sigma_X^2 = \mathrm{Var}(X)$, $\sigma_Y^2 = \mathrm{Var}(Y)$, $\sigma_{XY} = \mathrm{Cov}(X,Y)$

We define coefficient of correlation between X and Y as

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X\ \sigma_Y}.$$

<u>Theorem</u>: For any jointly distributed random variables (X, Y), $-1 \leq \rho_{X,Y} \leq 1$.

<u>Proof</u>: Let $U = \dfrac{X - \mu_X}{\sigma_X}$ and $V = \dfrac{Y - \mu_Y}{\sigma_Y}$

Then $E(U) = 0$, $E(V) = 0$

$E(U^2) = 1$, $E(V^2) = 1$

$$E(UV) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \rho_{X,Y}$$

Now consider the inequality

$$E(U-V)^2 \geq 0$$

$$\Rightarrow E(U^2) + E(V^2) - 2E(UV) \geq 0$$

$$\Rightarrow \quad E(UV) \leq 1 \qquad \cdots (1)$$

Similar consider the inequality

$$E(U+V)^2 \geq 0$$

$$\Rightarrow E(U^2) + E(V^2) + 2E(UV) \geq 0$$

$$\Rightarrow \quad E(UV) \geq -1 \qquad \cdots (2)$$

Combining (1) and (2) we get

$$-1 \leq \rho_{x,y} \leq 1$$

If $E(U-V)^2 = 0 \Rightarrow \rho_{x,y} = 1$

$\downarrow$

$\Rightarrow \quad P(U=V) = 1$

$\Rightarrow \quad P(X = cY + d, \quad c > 0) = 1$

ie X and Y are linearly related in a positive direction

$$E(U+V)^2 = 0 \rightarrow \rho_{X,Y} = -1$$

$$\downarrow$$

$$P(U = -V) = 1$$

$$\Rightarrow P(X = cY + d, \ c<0) = 1$$

ie X and Y are linearly related in a negative direction

Thus $\rho_{X,Y}$ is a measure of linear relationship between r.v's X and Y.

If $\rho_{X,Y} = 0$ we say that X and Y are linearly uncorrelated.

Theorem: If X and Y are independent then $\rho_{X,Y} = 0$. But the converse is not true.

Example:

$X \backslash Y$	-1	0	1	$p_X(x)$
0	0	$1/3$	0	$1/3$
1	$1/3$	0	$1/3$	$2/3$
$p_Y(y)$	$1/3$	$1/3$	$1/3$	

$$E(X) = 0\cdot\frac{1}{3} + 1\cdot\frac{2}{3} = 2/3$$

$$E(Y) = -1\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{3} = 0$$

$$E(XY) = 0\cdot(-1)\cdot 0 + 0\cdot 0\cdot\frac{1}{3} + 0(1)\cdot 0$$
$$+ 1\cdot(-1)\cdot\frac{1}{3} + 1\cdot 0\cdot 0 + 1\cdot 1\cdot\frac{1}{3} = 0$$

So $\operatorname{Cov}(X,Y) = 0 \Rightarrow \rho_{X,Y} = 0$

But X and Y are not independent here.